南京航空航天大學
NANJING UNIVERSITY OF AERONAUTICS AND ASTRONAUTICS

FixBi: Bridging Domain Spaces for Unsupervised Domain Adaptation

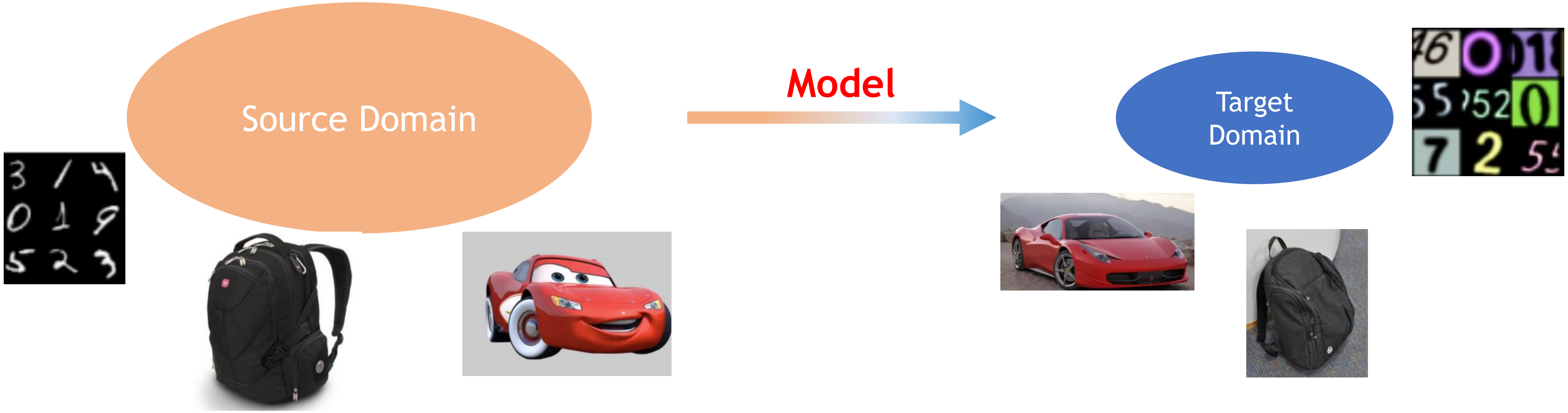
Jaemin Na[1], Heechul Jung[2], Hyung Jin Chang[3], and Wonjun Hwang[1]
[1]Ajou University, Korea, [2]Kyungpook National University, Korea, [3]University of Birmingham, UK
osial46@ajou.ac.kr, heechul@knu.ac.kr, h.j.chang@bham.ac.uk, wjhwang@ajou.ac.kr

CVPR 2021

Background

Unsupervised Domain Adaptation

UDA (Unsupervised Domain Adaptation) aims to address the problem of classifying unlabeled samples from the target domain whilst labeled samples are only available from the source domain and the data distributions are different in these two domains.



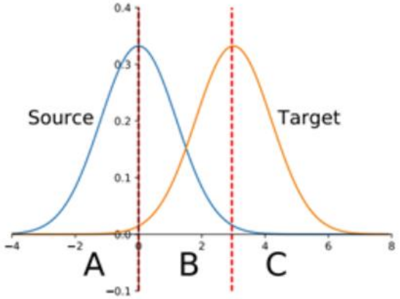
Source Domain $\sim P_S(X,Y)$ $\neq$ Target Domain $\sim P_T(Z,H)$

lots of **labeled** data

unlabeled or limited labels

$$D_S = \{(\mathbf{x}_i, y_i), \forall i \in \{1, \ldots, N\}\}$$

$$D_T = \{(\mathbf{z}_j, \mathbf{?}), \forall j \in \{1, \ldots, M\}\}$$

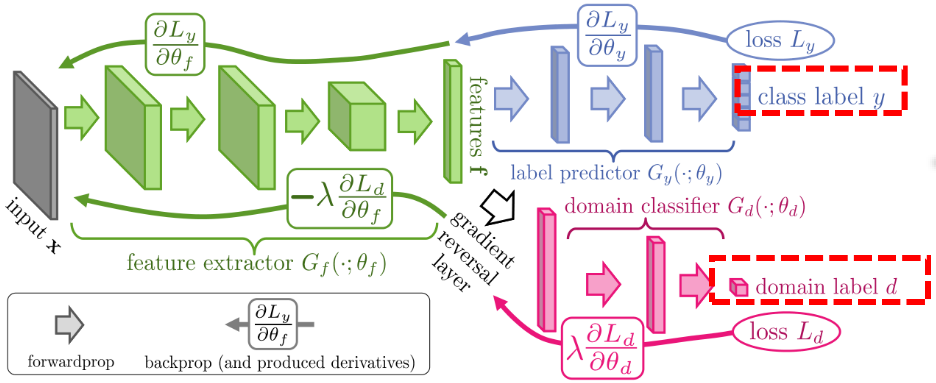


Background

Domain Alignment

- **DANN**

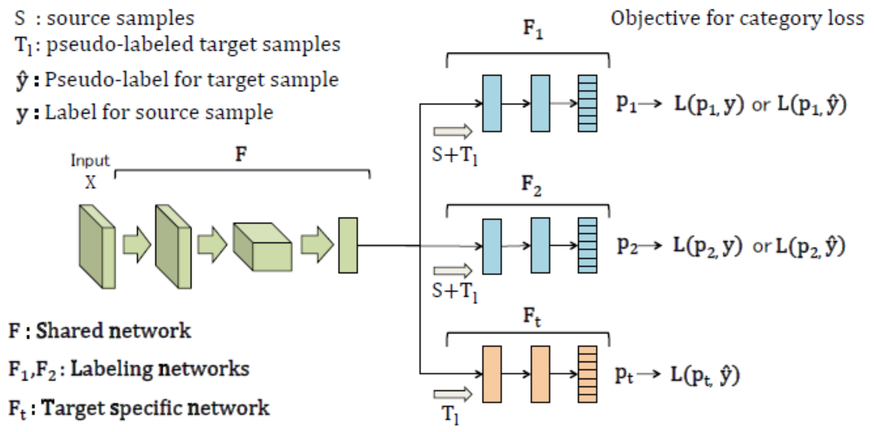
通过对抗，发现源域和目标域之间可迁移的**不变特征**（transferable features）



Domain-invariance: 面对这些特征，无法区分是来自源域还是目标域

Discriminativeness ：利用这些特征，可以很好地完成分类任务

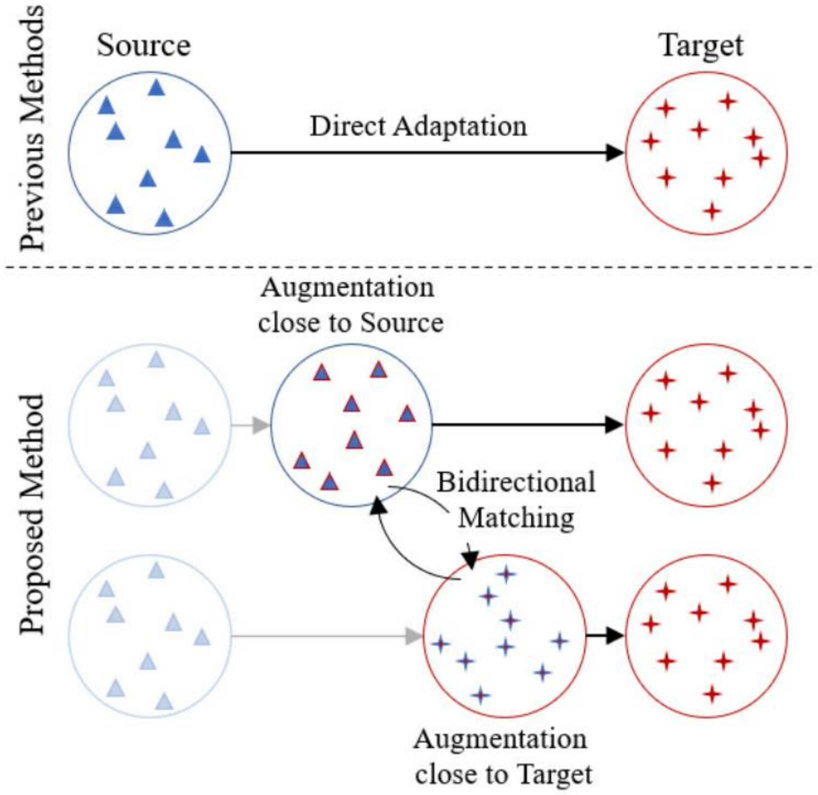
- **Asymmetric Tri-training for Unsupervised Domain Adaptation**



Labeling Network

Target Specific Network

Motivation



- **Top**: Previous methods try to adapt directly without any consideration of large domain discrepancies as there is large gap between source and target domain.

- **Bottom**: The methods in this paper try to utilize augmented domain between source and target domain for efficient domain adaptation.

FixBi - Fixed Ratio-based Mixup(warm up)

- This paper propose to use two fixed mixup ratios λ_{sd} and λ_{td} to make augmented samples.

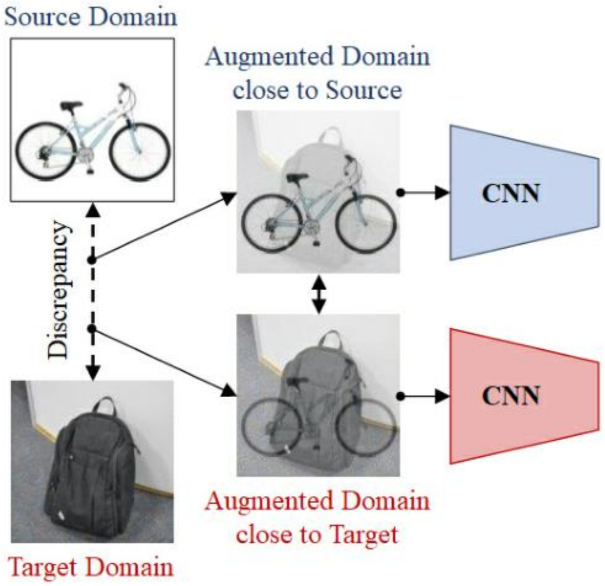
$$\tilde{x}_i^{st} = \lambda x_i^s + (1-\lambda) x_i^t$$
$$\tilde{y}_i^{st} = \lambda y_i^s + (1-\lambda) \hat{y}_i^t, \quad (1)$$

$$\lambda \in \{\lambda_{sd}, \lambda_{td}\}$$
$$s.t.\ \lambda_{sd} + \lambda_{td} = 1$$

"source-dominant model" (SDM)

"target-dominant model" (TDM)

- Two fixed mixup ratios-based models

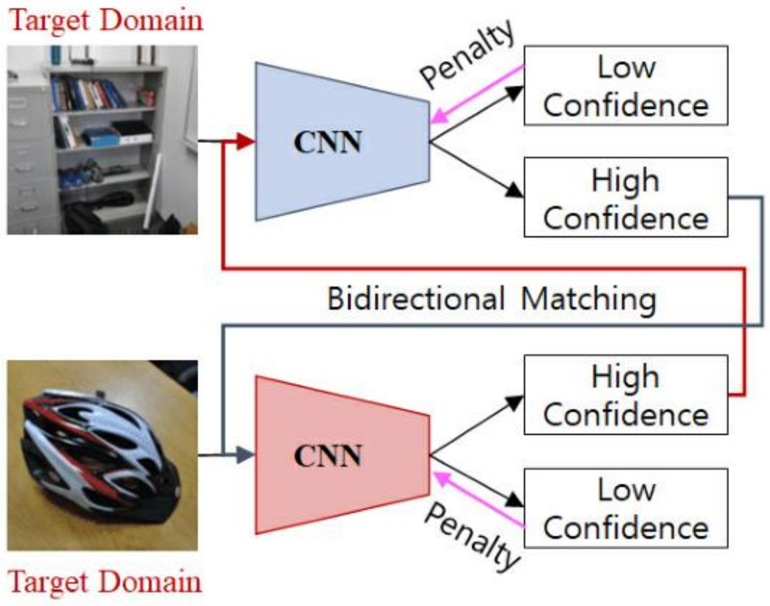


(a) Fixed Ratio-based Mixup

$$\mathcal{L}_{fm} = \frac{1}{B} \sum_{i=1}^{B} \hat{y}_i^{st} \log(p(y|\tilde{x}_i^{st})), \quad (2)$$

Ratio-based Models | Confidence-based Learning | Consistency Regularization

FixBi - Confidence-based Learning



(b) Confidence-based Learning

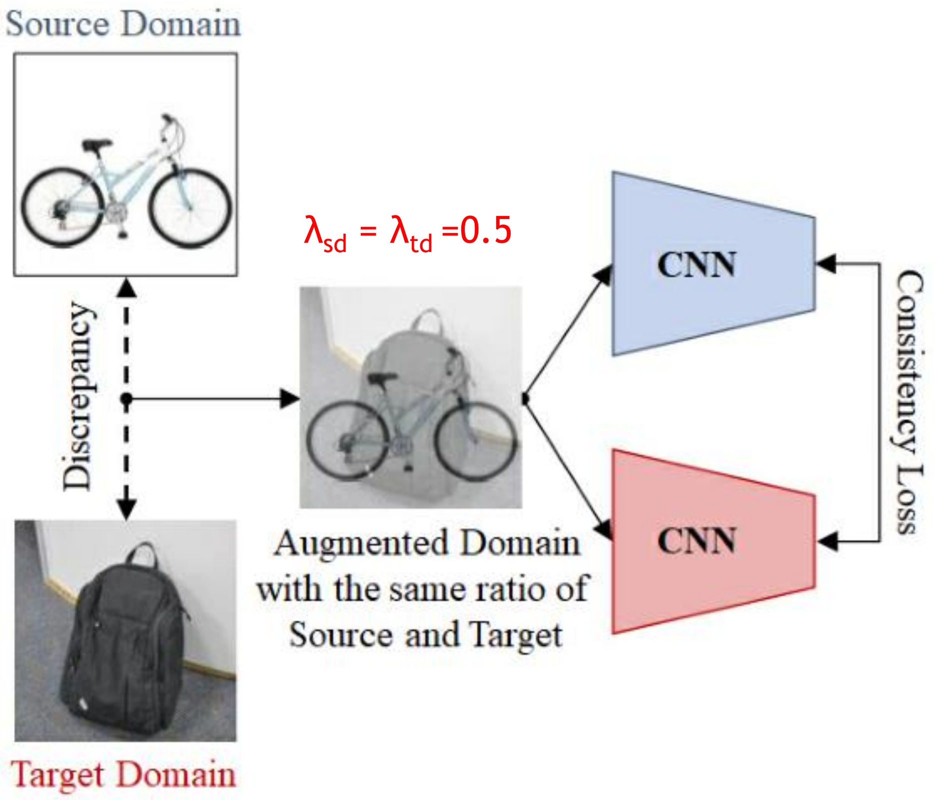
$$\underbrace{\quad}_{}^{mean - 2 \times std}$$

Loss : $$\mathcal{L}_{bim} = \frac{1}{B}\sum_{i=1}^{B} \mathbb{1}(max(p(y|x_i^t) > \tau)\hat{y}_i^t \log(q(y|x_i^t)), \quad (3)$$

- Pseudo-labels
  - Positive pseudo-labels for Bidirectional Matching
  - Negative pseudo-labels for self-penalization

Loss : $$\mathcal{L}_{sp} = \frac{1}{B}\sum_{i=1}^{B} \mathbb{1}(max(p(y|x_i^t) < \tau)\hat{y}_i^t \log(1 - p(y|x_i^t)). \quad (4)$$

FixBi - Consistency Regularization



(c) Consistency Regularization

$$\mathcal{L}_{cr} = \frac{1}{B}\sum_{i=1}^{B} \| p(y|\tilde{x}_i^{st}) - q(y|\tilde{x}_i^{st}) \|_2^2. \tag{5}$$

Ratio-based Models | Confidence-based Learning | Consistency Regularization

FixBi - Algorithm

Algorithm 1: FixBi Training Procedure

Input : Network weights w_{sd} and w_{td}, total epochs E, mini batch B, warm-up epochs k, mixup-ratios λ_{sd}, λ_{td}, and $\lambda_{cr}(=0.5)$, source samples x^s, target samples x^t, and mixup samples $\tilde{M}$.

for e=1 **to** E **do**
  for i=1 **to** B **do**
    Obtain $\tilde{M}_{sd}$ using Eq. (1) with λ_{sd};
    Obtain $\tilde{M}_{td}$ using Eq. (1) with λ_{td};
    Update $\boxed{\mathcal{L}_{fm}(\tilde{M}_{sd}; w_{sd})}$ and $\boxed{\mathcal{L}_{sp}(x^t; w_{sd})}$;
    Update $\boxed{\mathcal{L}_{fm}(\tilde{M}_{td}; w_{td})}$ and $\boxed{\mathcal{L}_{sp}(x^t; w_{td})}$;
    if $e > k$ **then** Warm up = 100
      if $max(y|x^t; w_{td}) > \tau_{td}$ **then**
        Update $\mathcal{L}_{bim}(x^t; w_{sd})$;
      end 这个T是怎么改变的?
      if $max(y|x^t; w_{sd}) > \tau_{sd}$ **then**
        Update $\mathcal{L}_{bim}(x^t; w_{td})$;
      end
      Obtain $\tilde{M}_{cr}$ using Eq. (1) with λ_{cr};
      Update $\mathcal{L}_{cr}(\tilde{M}_{cr}; w_{sd})$;
      Update $\mathcal{L}_{cr}(\tilde{M}_{cr}; w_{td})$;
    end
  end
end

Output: Learned model parameters w_{sd} and w_{td}.

Experiments - SOTA

Table 4. Accuracy (%) on Office-31 for unsupervised domain adaptation (ResNet-50). The best accuracy is indicated in bold and the second best one is underlined. * Reproduced by [5]

Method	A→W	D→W	W→D	A→D	D→A	W→A	Avg
ResNet-50 [13]	68.4±0.2	96.7±0.1	99.3±0.1	68.9±0.2	62.5±0.3	60.7±0.3	76.1
DANN [8]	82.0±0.4	96.9±0.2	99.1±0.1	79.7±0.4	68.2±0.4	67.4±0.5	82.2
MSTN* [42]	91.3	98.9	**100.0**	90.4	72.7	65.6	86.5
CDAN+E [23]	94.1±0.1	98.6±0.1	**100.0±0.0**	92.9±0.2	71.0±0.3	69.3±0.3	87.7
DMRL [41]	90.8±0.3	99.0±0.2	**100.0±0.0**	93.4±0.5	73.0±0.3	71.2±0.3	87.9
SymNets [47]	90.8±0.1	98.8±0.3	**100.0±0.0**	93.9±0.5	74.6±0.6	72.5±0.5	88.4
GSDA [16]	95.7	99.1	**100**	94.8	73.5	74.9	89.7
CAN [17]	94.5±0.3	99.1±0.2	<u>99.8±0.2</u>	95.0±0.3	<u>78.0±0.3</u>	77.0±0.3	90.6
SRDC [36]	<u>95.7±0.2</u>	<u>99.2±0.1</u>	**100.0±0.0**	**95.8±0.2**	76.7±0.3	77.1±0.1	90.8
RSDA-MSTN [11]	**96.1±0.2**	**99.3±0.2**	**100.0±0.0**	<u>95.8±0.3</u>	77.4±0.8	<u>78.9±0.3</u>	<u>91.1</u>
FixBi (Ours)	**96.1±0.2**	**99.3±0.2**	**100.0±0.0**	95.0±0.4	**78.7±0.5**	**79.4±0.3**	**91.4**

Experiments - SOTA

Table 5. Accuracy (%) on Office-Home for unsupervised domain adaptation (ResNet-50). The best accuracy is indicated in bold and the second best one is underlined. * Reproduced by [11]

Method	Ar→Cl	Ar→Pr	Ar→Rw	Cl→Ar	Cl→Pr	Cl→Rw	Pr→Ar	Pr→Cl	Pr→Rw	Rw→Ar	Rw→Cl	Rw→Pr	Avg
ResNet-50 [13]	34.9	50	58	37.4	41.9	46.2	38.5	31.2	60.4	53.9	41.2	59.9	46.1
DANN [8]	45.6	59.3	70.1	47	58.5	60.9	46.1	43.7	68.5	63.2	51.8	76.8	57.6
CDAN [23]	49	69.3	74.5	54.4	66	68.4	55.6	48.3	75.9	68.4	55.4	80.5	63.8
MSTN* [42]	49.8	70.3	76.3	60.4	68.5	69.6	61.4	48.9	75.7	70.9	55	81.1	65.7
SymNets [47]	47.7	72.9	78.5	64.2	71.3	74.2	63.6	47.6	79.4	73.8	50.8	82.6	67.2
GSDA [16]	**61.3**	76.1	79.4	65.4	73.3	74.3	65	53.2	80	72.2	$\underline{60.6}$	83.1	70.3
GVB-GD [7]	57	74.7	79.8	64.6	74.1	74.6	65.2	$\underline{55.1}$	81	74.6	59.7	84.3	70.4
RSDA-MSTN [11]	53.2	**77.7**	**81.3**	66.4	74	76.5	$\underline{67.9}$	53	**82**	75.8	57.8	$\underline{85.4}$	70.9
SRDC [36]	52.3	76.3	$\underline{81}$	**69.5**	$\underline{76.2}$	$\underline{78}$	**68.7**	53.8	$\underline{81.7}$	$\underline{76.3}$	57.1	85	$\underline{71.3}$
FixBi (Ours)	$\underline{58.1}$	$\underline{77.3}$	80.4	$\underline{67.7}$	**79.5**	**78.1**	65.8	**57.9**	$\underline{81.7}$	**76.4**	**62.9**	**86.7**	**72.7**

Table 6. Accuracy (%) on VisDA-2017 for unsupervised domain adaptation (ResNet-101). The best accuracy is indicated in bold and the second best one is underlined. * Reproduced by [5]

Method	aero	bicycle	bus	car	horse	knife	motor	person	plant	skate	train	truck	Avg
ResNet-101 [13]	72.3	6.1	63.4	91.7	52.7	7.9	80.1	5.6	90.1	18.5	78.1	25.9	49.4
DANN [8]	81.9	77.7	82.8	44.3	81.2	29.5	65.1	28.6	51.9	54.6	82.8	7.8	57.4
DAN [22]	68.1	15.4	76.5	$\underline{87}$	71.1	48.9	82.3	51.5	88.7	33.2	$\underline{88.9}$	42.2	61.1
MSTN* [42]	89.3	49.5	74.3	67.6	90.1	16.6	**93.6**	70.1	86.5	40.4	83.2	18.5	65.0
JAN [24]	75.7	18.7	82.3	86.3	70.2	56.9	80.5	53.8	92.5	32.2	84.5	$\underline{54.5}$	65.7
DM-ADA [43]	-	-	-	-	-	-	-	-	-	-	-	-	75.6
DMRL [41]	-	-	-	-	-	-	-	-	-	-	-	-	75.5
MODEL [20]	94.8	73.4	68.8	74.8	93.1	$\underline{95.4}$	88.6	$\underline{84.7}$	89.1	84.7	83.5	48.1	81.6
STAR [25]	95	84	$\underline{84.6}$	73	91.6	91.8	85.9	78.4	94.4	84.7	87	42.2	$\underline{82.7}$
CAN [17]	**97**	$\underline{87.2}$	82.5	74.3	**97.8**	**96.2**	90.8	80.7	$\underline{96.6}$	**96.3**	87.5	**59.9**	**87.2**
FixBi (Ours)	$\underline{96.1}$	**87.8**	**90.5**	**90.3**	$\underline{96.8}$	95.3	$\underline{92.8}$	**88.7**	**97.2**	$\underline{94.2}$	**90.9**	25.7	**87.2**

FixBi - Experiments

- Comparison of different mixup-ratio rules

Table 1. Comparison of three different mixup-ratio rules on the task A→W.

Type	w/o $\mathcal{L}_{bim}$		w/ $\mathcal{L}_{bim}$	
	SDM	TDM	SDM	TDM
Random	86.5±1.0	85.3±0.9	86.7±0.8	85.6±0.7
Range	86.0±1.7	29.6±6.8	83.3±6.2	81.0±5.4
Fixed (Ours)	86.3±0.6	86.0±0.7	89.3±0.4	90.1±0.3

The ratio randomly sampled from the beta distribution

Office-31 A -> W

FixBi - Experiments

☐ Why is better to use a fixed ratio ?

The two networks have two different perspectives through fixed ratio-based mixup.

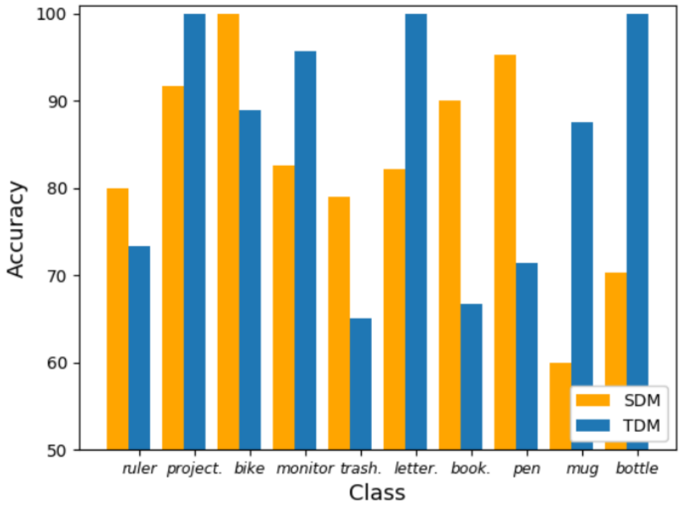


Figure 3. Class-wise accuracy (%) on the task A→W of the Office-31. Best viewed in color.

Table 2. Comparison of ensemble networks on Office-31.

Method	$(\lambda_{sd}, \lambda_{td})$	A→W	D→W	W→D	A→D	D→A	W→A	Avg
Single-perspective	(0.3, 0.3)	88.6	96.5	100.0	85.6	69.4	65.1	84.2
	(0.7, 0.7)	89.2	96.5	100.0	85.5	69.1	67.8	84.7
Two-perspective (Ours)	(0.7, 0.3)	**90.1**	**98.5**	**100.0**	**88.4**	**72.5**	**72.5**	**87.0**

Experiments - Ablation study

Table 3. Ablation results (%) of investigating the effects of our components on Office-31.

$\mathcal{L}_{DANN}$	$\mathcal{L}_{fm}$	$\mathcal{L}_{bim}$	$\mathcal{L}_{sp}$	$\mathcal{L}_{cr}$	A→W	D→W	W→D	A→D	D→A	W→A	Avg
✓					82.0	96.9	99.1	79.7	68.2	67.4	82.2
✓	✓				86.5	98.4	**100.0**	85.5	71.4	71.5	85.5
✓	✓	✓			90.1	98.5	**100.0**	88.4	72.5	72.5	87.0
✓	✓	✓	✓		92.3	98.6	**100.0**	90.4	76.3	74.1	88.6
✓	✓	✓	✓	✓	**94.2**	**99.3**	**100.0**	**91.3**	**76.5**	**74.3**	**89.3**

Thanks